

Supplementary Table 1 - MuTect CCG>CAG filter training set details

Sample	Library Date	n	Ntot	Npass	Nreject	FP	FN	TP	TN	EFP	EFN	spec	sens	PPV	NPV	FNR	FNR CI 95%	FPR	FPR CI 95%	FoxoG
Bladder tumor 1	11/17/2011	1	450	391	59	1.2	-1.1	57.8	392.1	0.30	5.01	0.9971	1.0202	0.9804	1.0029	-0.3%	2.3%	0.3%	0.5%	12.6%
Bladder tumor 2	11/17/2011	1	1400	434	966	1.8	5.0	964.2	429.0	0.37	6.86	0.9958	0.9948	0.9981	0.9885	1.2%	4.3%	0.4%	0.7%	69.2%
Bladder tumor 3	11/17/2011	1	1673	1228	445	3.3	-12.0	441.7	1240.0	0.42	9.90	0.9974	1.0279	0.9927	1.0098	-1.0%	0.6%	0.3%	0.4%	25.7%
Cervical tumor 1	11/18/2011	1	744	44	700	0.0	1.0	700.0	43.0	0.00	1.00	1.0000	0.9986	1.0000	0.9773	2.3%	6.8%	0.0%	0.0%	94.2%
Cervical tumor 2	11/18/2011	1	429	305	124	0.9	-3.1	123.1	308.1	0.28	3.62	0.9971	1.0256	0.9928	1.0101	-1.0%	1.4%	0.3%	0.6%	28.0%
Cervical tumor 3	12/19/2011	1	157	144	13	0.8	1.6	12.3	142.4	0.27	3.80	0.9948	0.8869	0.9423	0.9891	1.1%	6.4%	0.5%	1.1%	8.8%
Head & neck tumor 1	2/15/2011	1	258	252	6	1.3	-8.4	4.7	260.4	0.31	4.73	0.9949	-1.2479	0.7758	1.0333	-3.3%	0.4%	0.5%	0.9%	-1.4%
Head & neck tumor 2	5/31/2011	1	5143	205	4938	3.5	2.6	4934.5	202.4	0.62	6.74	0.9832	0.9995	0.9993	0.9874	1.3%	7.8%	1.7%	2.6%	96.0%
Head & neck tumor 3	9/29/2011	1	168	135	33	1.1	-2.2	31.9	137.2	0.35	3.77	0.9921	1.0753	0.9669	1.0166	-1.7%	3.9%	0.8%	1.5%	17.7%
Lung Squamous tumor 1	8/8/2011	1	1142	424	718	3.5	0.1	714.5	423.9	0.50	9.16	0.9917	0.9999	0.9951	0.9999	0.0%	4.3%	0.8%	1.2%	62.6%
Lung Squamous tumor 2	8/8/2011	1	1268	431	837	4.8	4.9	832.2	426.1	0.65	9.75	0.9889	0.9941	0.9943	0.9886	1.1%	5.7%	1.1%	1.6%	66.0%
Lung Squamous tumor 3	8/8/2011	1	1432	388	1044	3.9	-5.3	1040.1	393.3	0.55	8.91	0.9902	1.0052	0.9963	1.0137	-1.4%	3.2%	1.0%	1.4%	72.3%
Rhabdoid tumor 1	5/16/2011	1	20	16	4	0.0	0.0	4.0	16.0	0.00	0.00	1.0000	1.0000	1.0000	1.0000	0.0%	0.0%	0.0%	0.0%	20.0%
Rhabdoid tumor 2	9/29/2011	1	442	11	431	0.4	-0.5	430.6	11.5	0.40	0.71	0.9664	1.0012	0.9991	1.0455	-4.5%	8.3%	3.4%	13.4%	97.3%
Neuroblastoma tumor 1	1/9/2012	1	32	25	7	0.0	1.0	7.0	24.0	0.00	1.00	1.0000	0.8750	1.0000	0.9600	4.0%	12.0%	0.0%	0.0%	25.0%
Neuroblastoma tumor 2	1/9/2012	1	6645	35	6610	2.5	10.5	6607.5	24.5	1.12	3.39	0.9074	0.9984	0.9996	0.7000	30.0%	49.4%	9.3%	21.7%	99.6%
Neuroblastoma tumor 3	1/9/2012	1	2204	24	2180	0.2	4.4	2179.8	19.6	0.17	2.36	0.9915	0.9980	0.9999	0.8148	18.5%	38.2%	0.8%	3.4%	99.1%
Neuroblastoma tumor 4	1/9/2012	1	6094	55	6039	1.1	5.4	6037.9	49.6	0.64	2.57	0.9781	0.9991	0.9998	0.9016	9.8%	19.2%	2.2%	6.0%	99.2%
Neuroblastoma tumor 5	1/6/2012	1	19	15	4	0.1	-0.5	3.9	15.5	0.14	0.71	0.9909	1.1489	0.9643	1.0333	-3.3%	6.1%	0.9%	3.7%	17.7%
Thyroid Tumor 1	12/6/2011	1	169	26	143	0.0	2.0	143.0	24.0	0.00	1.41	1.0000	0.9862	1.0000	0.9231	7.7%	18.6%	0.0%	0.0%	85.8%
Thyroid Tumor 2	12/16/2011	1	90	28	62	1.6	-0.4	60.4	28.4	0.78	1.55	0.9478	1.0065	0.9748	1.0140	-1.4%	9.6%	5.2%	13.1%	66.7%
Thyroid Tumor 3	12/16/2011	1	1164	19	1145	0.5	0.8	1144.5	18.2	0.47	1.08	0.9748	0.9993	0.9996	0.9567	4.3%	15.8%	2.5%	10.1%	98.4%
Thyroid Tumor 4	5/16/2011	1	21	18	3	0.6	-1.9	2.4	19.9	0.32	1.37	0.9725	4.3333	0.8125	1.1042	-10.4%	4.8%	2.8%	7.5%	2.7%
Thyroid Tumor 5	5/16/2011	1	2006	49	1957	1.1	2.2	1955.9	46.8	0.45	3.13	0.9769	0.9989	0.9994	0.9549	4.5%	17.3%	2.3%	5.1%	97.6%
Thyroid Tumor 6	5/16/2011	1	75	29	46	0.2	0.5	45.8	28.5	0.19	1.23	0.9935	0.9896	0.9960	0.9834	1.7%	10.2%	0.6%	2.6%	61.7%
Thyroid Tumor 7	8/11/2011	1	73	28	45	0.7	1.2	44.3	26.8	0.47	1.68	0.9758	0.9744	0.9852	0.9583	4.2%	16.2%	2.4%	7.6%	62.3%
Thyroid Tumor 8	2/15/2011	1	36	33	3	0.0	-0.7	3.0	33.7	0.03	0.83	0.9991	1.3014	0.9896	1.0208	-2.1%	2.9%	0.1%	0.4%	6.3%
Thyroid Tumor 9	2/15/2011	1	37	33	4	0.2	-0.6	3.8	33.6	0.19	0.79	0.9945	1.1961	0.9531	1.0189	-1.9%	2.9%	0.6%	2.2%	8.6%
Thyroid Tumor 10	2/15/2011	1	26	25	1	0.2	-0.6	0.8	25.6	0.21	0.76	0.9919	3.8000	0.7917	1.0233	-2.3%	3.8%	0.8%	3.2%	0.8%
Thyroid Tumor 11	8/11/2011	1	481	14	467	0.0	0.0	467.0	14.0	0.00	0.00	1.0000	1.0000	1.0000	1.0000	0.0%	0.0%	0.0%	0.0%	97.1%
Thyroid Tumor 12	12/16/2011	1	441	29	412	0.4	0.6	411.6	28.4	0.21	1.85	0.9875	0.9986	0.9991	0.9807	1.9%	14.7%	1.3%	3.4%	93.5%
TOTALS		31	34339	4893	29446	67.5	6.1	29378.5	4886.9	3.55	24.33	0.9864	0.9998	0.9977	0.9988	0.1%	1.6%	1.4%	1.6%	85.6%

Key:	Definition
sample	sample id
Ntot	total number of raw mutations called by MuTect
Npass	number of mutations passing OxoG filter
Nreject	number of mutations rejected by OxoG filter
FP	estimated number of false positive mutations based on null model (OxoG artifact mutations passed by filter)
FN	estimated number of false negative mutations based on null model (non-OxoG artifact mutations rejected by filter)
TP	estimated number of true positive mutations based on null model (OxoG artifact mutations rejected by filter)
TN	estimated number of true negative mutations based on null model (non-OxoG artifact mutations passed by filter)
EFP	Standard error on FP estimate
EFN	Standard error on FN estimate
spec	specificity: $TN/(TN+FP)$
sens	sensitivity: $TP/(TP+FN)$
PPV	$TP/(TP+FP)$ positive predictive value
NPV	$TN/(TN+FN)$ negative predictive value
FNR	$FN/(TN+FN)=1-NPV$ false negative rate
FNR CI 95%	FNR upper confidence limit 95%
FPR	$FP/(TP+FP)=1-PPV$ false positive rate
FPR CI 95%	FPR upper confidence limit 95%
FoxoG	proportion of artifact in raw MuTect calls

Note:

Estimates of FP and FN are based on projections of the null model (non-C>A,G>T mutations) onto the 2d histogram of the artifact mode (C>A,G>T) mutations in FoxoG, Tumor_id space. Inherent statistical fluctuations limit the accuracy of the estimate and can even result in FN<0 estimates. The scale of the statistical uncertainty is quantified by the EFP and EFN values which is based on the binomial standard error estimate for FP and FN.